

Model Validation: A Connection Between Robust Control and Identification

Roy S. Smith, *Member, IEEE*, and John C. Doyle

Abstract—Modern robust control synthesis techniques aim at providing robustness with respect to uncertainty in the form of both additive noise and plant perturbations. On the other hand, the most popular system identification methods assume that all uncertainty is in the form of additive noise. This has hampered the application of robust control methods to practical problems. This paper begins to address the gap between the models used in control synthesis and those obtained from identification experiments by considering the connection between uncertain models and data. The model validation problem addressed here is: given experimental data and a model with both additive noise and norm-bounded perturbations, is it possible that the model could produce the observed input-output data? This problem is studied for the standard H_∞/μ framework models. A necessary condition for such a model to describe an experimental datum is obtained. Furthermore, for a large class of models, in the robust control framework, this condition is computable as the solution of a quadratic optimization problem.

I. INTRODUCTION

ROBUST control theory now gives the engineer the power to describe physical systems with a model which includes a particular type of uncertainty: block structured, norm-bounded uncertainty entering the model in a linear fractional manner. Linear models in which the only uncertainty is in the form of additive noise cannot account for a loss of stability not predicted by the nominal model. Robust control models can capture this feature, essentially being able to include unmodeled but bounded dynamics.

Once a system is modeled, and the engineer is confident of the applicability of the model, the theory gives techniques for designing closed-loop systems which are theoretically robust. The theory makes no statements about the performance or the stability of the actual physical system. Therein lies a problem for the engineer. Before the robust control methods can be applied, the uncertainty must, in some sense, be identified. Current identification methods are well developed in the case where all of the residuals, or uncertainty, are attributed to additive noise. For models with both additive noise and norm-bounded perturbations, no such identification methods exist.

Manuscript received January 21, 1991; revised November 21, 1991. Paper recommended by Associate Editor at Large, M. P. Polis. This work was supported by the National Aeronautics and Space Administration, The National Science Foundation, The Office of Naval Research, and by Rockwell International.

R. S. Smith is with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106.

J. C. Doyle is with the California Institute of Technology, Pasadena, CA 91125.

IEEE Log Number 9200619.

Once a model has been determined, perhaps by ad-hoc methods, there must be some method of evaluating its applicability to the actual physical system. The engineer must be confident that the model will describe all input-output behaviors of the system. This condition can never be guaranteed but it is possible to test a necessary condition: that the model be able to describe all observed input-output behaviors of the system. This is simply the model validation question to be discussed here.

A. Context

The model validation question posed above can be considered for any model of a physical system. This paper will restrict itself to models fitting the standard H_∞ and μ synthesis framework. To be directly applicable to a particular design framework, the assumptions of that framework also have to be applied to the model validation problem. The choice of design/analysis framework determines the properties of the model validation problem and the experiments which are applicable to the formal model validation study.

In particular, this paper will assume that all unknown signals (noise and disturbances) are signals of bounded energy. Furthermore, it is assumed that the measurements of the plant inputs and outputs are available in the frequency domain. These assumptions are very restrictive when applied to experiment design. Smith [1] does give an example of a useful experiment and the application of the techniques discussed here.

The model validation problem can be applied to other design paradigms. Poolla [2] is investigating its application to the discrete time, bounded energy case. The application to l_1 problems is also currently under investigation. The significance of this approach is that it is applied in the context of a design/analysis methodology. The fact that application to current, popular methodologies lead to a very restricted set of allowable experiments can be taken as an argument for developing design methodologies with more realistic assumptions on the signal sets.

Standard approaches to identification typically use stochastic assumptions on the noise. Furthermore, dynamic uncertainties are usually neglected. Ljung [3] discusses the standard identification approaches in detail. Identification in H_∞ is an area of growing research interest. The reader is referred to the works of Helmicki, Jacobson, and Nett [4] and [5], Gu and Khargonekar [6], Mäkilä [7], and Mäkilä and Partington [8]. Although these approaches are encouraging, the results to date do not result in models matching the H_∞/μ design methodology. In most of the work referenced above, only

additive, rather than fractional perturbations are considered and the assumptions on the noise do not match those used in the design methodology. More importantly, the underlying system is assumed to be linear and time-invariant. Some assumptions about the systems being identified are required in order to obtain results on the suitability of any identification approach.

The approach of this paper is significantly different in that no assumptions on the nature of the physical system are needed. Model validation is a test of a given model against experimental input-output data. The assumptions of the design framework apply to the model, not the physical system.

B. Organization of the Paper

The paper is organized as follows. Section II introduces robust control models. The assumptions most commonly used in the practical application of the H_∞/μ design methodology are presented. Model validation is discussed in detail in Section III. A generic model for identification and validation is introduced and the model validation problem is stated formally.

Section IV considers a constant matrix version of the model validation problem. This is shown to be strongly connected with the μ problem. The approach to μ developed by Fan and Tits [9] is applied to the model validation problem. A function ψ is defined and shown to apply to model validation in the same manner that μ applies to robust analysis.

Section V develops a "skewed" μ problem. A skewed version of the constant matrix model validation problem is also developed. This can be viewed as determining the minimum amount of additive noise, in the presence of unity bounded perturbations, required to describe the experimental observation. Skewed μ (or ψ) problems inherit the convexity properties of the nonskewed versions. The results of Section V are given without proof. For full details refer to [1].

In Section VI the original model validation problem presented in Section III is reconsidered. A necessary condition for the model to describe the experiment is given in terms of constant matrix, skewed, model validation problems at each frequency. Section VII concludes the paper with a summary of the results and suggests future research directions.

C. Notation

Capital letters will denote matrices or matrix valued functions. A^T is the transpose of A , and A^* is the complex conjugate transpose. $\bar{\sigma}(A)$ is the maximum singular value of A , $\dim(A)$ is the dimension of A , and $\ker(A)$ is the kernel of A . The convex hull of a set S is denoted by $\text{co}(S)$. The inner product of two vectors $x, y \in \mathbb{C}^n$, is denoted by $\langle x, y \rangle$, where $\langle x, y \rangle = x^*y$.

Block diagrams will be used to represent interconnections of systems. For example, the generic robust control model is shown in Fig. 1. This diagram represents the equations

$$\begin{aligned} z &= P_{11}v + P_{12}w \\ e &= P_{21}v + P_{22}w \\ v &= \Delta z. \end{aligned}$$

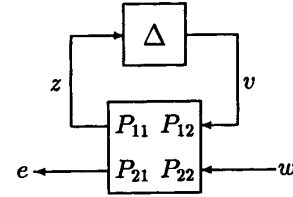


Fig. 1. Generic model structure including uncertainty.

In terms of the input w and output e , this can be expressed as a linear fractional transformation (LFT)

$$e = [P_{21}\Delta(I - P_{11}\Delta)^{-1}P_{12} + P_{22}]w = F_u(P, \Delta)w$$

where the subscript u indicates that Δ is used to close the upper loop.

II. ROBUST CONTROL MODELS

The framework presented here follows that of Doyle [10]. A detailed presentation of more general analysis results is given by Packard [11]. The set of uncertain models will be formalized with specific assumptions on $P(s)$, Δ , and the unknown signals w . Significantly more generality is possible but will complicate the exposition. Note that the assumptions given here are those used in the majority of the practical applications of H_∞ and μ synthesis to date. These assumptions also reduce analysis problems to constant matrix problems at each frequency.

Linear models, for example, $P(s)$ or simply P , are assumed to be in $\mathbb{R}H_\infty$, real rational functions with

$$\|P(s)\|_\infty = \sup_\omega \bar{\sigma}[P(j\omega)] < \infty.$$

Unknown signals w are assumed to be elements of BL_2 , where

$$BL_2 = \left\{ w \mid \|w\|_2^2 = \int_{-\infty}^{\infty} |w(t)|^2 dt < 1 \right\}.$$

For a vector valued w , $|w(t)|$ denotes the Euclidean norm of $w(t)$.

The generic structure of the robust control model is shown in Fig. 1. The perturbations Δ are assumed to be causal, stable, linear, time-invariant, and block diagonal, with m square blocks each of dimension k_i , $i = 1, \dots, m$. The assumption of square blocks is without loss of generality.

Define Δ as the set of all perturbations of the prescribed structure

$$\Delta = \{\text{diag}(\Delta_1 \cdots \Delta_m) \mid \dim(\Delta_i) = k_i \times k_i\}.$$

The unit ball of Δ is given by

$$B\Delta = \{\Delta \mid \Delta \in \Delta, \|\Delta\|_\infty \leq 1\}.$$

At each frequency, ω , $\Delta_i(j\omega) \in \mathbb{C}^{k_i \times k_i}$, and $\bar{\sigma}(\Delta(j\omega)) \leq 1$.

The robust control model is represented by

$$e = F_u(P, \Delta)w, \quad \Delta \in B\Delta.$$

The model is robustly stable iff it is nominally stable and

$$\sup_{\omega} \mu(P(j\omega)) < 1.$$

A formal definition of μ is given in (6).

III. MODEL VALIDATION

An engineer is faced with the problem of selecting good models for the analysis of a system and the synthesis of controllers. Robust control models are more complex in that uncertainty can now enter in a fractional manner and bounds on that uncertainty must be chosen. No theory for systematically doing this exists. It is in fact a poorly posed problem; the physical system can only be observed by input-output measurements and, modulo considerations of observability from a particular output the residuals can be attributed either entirely to additive noise or entirely to norm-bounded perturbations. In practice, an engineer will run many experiments to attempt to isolate the effects of noise from those of perturbations.

An assumption, inherent in the use of any model, is that the model can describe any observed input-output behavior of the physical system. It is not possible to test this, however, it is possible to find a necessary condition. This is just the model validation question: Can the model account for all of the previously observed input-output behavior? This will be formulated more rigorously in the context of robust control models in Section III-B.

In the case where the residuals are attributed entirely to additive noise, identification procedures can give a noise model which guarantees that the model is consistent with all past data. For robust control models no such methods exist. For these models, validation techniques are required to test the suitability of the model.

The model validation theory described here has two additional uses. Large systems of many interconnected uncertain components can lead to models with many perturbation blocks. For simplicity in the design process, an engineer may wish to reduce the complexity of the interconnection structure by aggregating several perturbation blocks into a single "covering" block. Model validation gives a means of testing these reduced models against experimental data from the physical system.

Of significant practical engineering interest is the problem of fault detection. Given a design model and a controller in operation, the model validation theory gives a means of continuously assessing whether or not the physical system is still described by the design model. It will be seen that the techniques presented here produce the perturbation and noise that come closest to satisfying the conditions of the model. Gradual deterioration in a system may manifest itself as increasing perturbations and/or noise required for accountability of the data. A sudden failure may be identified by a sudden jump in the size of the required perturbations and/or noise.

A. A Generic Model for Identification and Validation

For the purposes of discussing identification and model validation, the generic structure, illustrated in Fig. 1 is

refined. Fig. 2 shows the structure that will be used throughout as the generic identification and model validation structure. In identification experiments certain inputs to the system are known. These are partitioned from the remaining, unknown inputs and denoted by u . As in the previous sections, w represents the unknown inputs from a specified set. The output y represents the measured outputs and is now also assumed to be known.

B. Formulation of the Model Validation Problem

It is assumed that the model P and block structure Δ are given, perhaps by ad-hoc identification methods, or first-principles modeling. In practice, a series of experiments would have been run. The problem posed here is with respect to a single experiment. In general, one would apply these methods to all of the experimental data.

Referring to the model structure of Fig. 2, the model validation problem can be formulated as follows.

Problem 3.1 (Model Validation): Given a model P , with

$$\sup_{\omega} \mu(P_{11}(j\omega)) < 1$$

and an experimental datum (u, y) , does there exist (w, Δ) , $w \in BL_2$, $\Delta \in B\Delta$, such that

$$y = F_u(P, \Delta) \begin{bmatrix} w \\ u \end{bmatrix}. \quad (1)$$

This simply asks the question "is there an element of the model set and an element of the unknown input signal set such that the observed datum is produced exactly?" Any (w, Δ) pair, with Δ having the correct structure ($\Delta \in B\Delta$), meeting the equality constraint (1) will be referred to as feasible. If a feasible (w, Δ) also satisfies $w \in BL_2$ and $\|\Delta\|_{\infty} \leq 1$, it is referred to as admissible. The model validation question is simply, does there exist an admissible (w, Δ) ? The approach taken will be to find the smallest feasible (w, Δ) .

Note that no statement is made relating the particular element of the model set or the particular element of the input signal set to any physical system or signal. Such a relationship does not exist. If no (w, Δ) pair meeting the above requirements exists, then the model cannot account for the observed behavior. Such a tool is of use in culling inappropriate models from a group of candidate models.

The model validation test is therefore a necessary condition for any model to describe a physical system. Model validation is a misleading term; strictly speaking it is never possible to validate a model, only to invalidate it. The fact that every experiment can be accounted for in this manner provides some confidence, but little hard information about the applicability of the model. There may be experiments, as yet unperformed, which will invalidate the model.

For those concerned with the requirement that the datum be produced exactly, note that physically realistic models will have some weighted part of w adding to y thereby modeling the effect of measurement noise. Note also that by assumption, the problem is restricted to the consideration of models which are robustly stable. This is a reasonable assumption from an engineering point of view, but will be relaxed in Section IV-A.

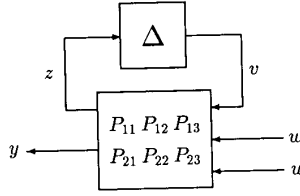


Fig. 2. The generic structure for identification and model validation problems.

IV. MODEL VALIDATION—CONSTANT MATRIX CASE

The model validation problem is now considered for the constant matrix case. In other words, P and Δ are complex matrices and y , u , and w are complex valued vectors. $\|w\|$ denotes the Euclidean norm and $\Delta \in B\Delta$ implies that Δ has the appropriate structure and $\bar{\sigma}(\Delta) \leq 1$. Section VI will show that the solution of Problem 3.1 is obtained in terms of the constant matrix problems studied here, and in Section V.

A constant matrix model validation problem is now posed. This is slightly more general than the constant matrix version of Problem 3.1 in that the assumption about the robust stability of P has been removed.

Problem 4.1 (Model Validation—Constant Matrix): Given a model P and an experimental datum (u, y) , what is the smallest $\|w\|$ and $\|\Delta\|$, $\Delta \in B\Delta$ such that

$$y = F_u(P, \Delta) \begin{bmatrix} w \\ u \end{bmatrix}.$$

The following assumptions are introduced and applied to the problem.

Assumption A0: $y - P_{23}u \neq 0$, $y - P_{23}u \in \text{Range}[P_{21} P_{22}]$ and $\dim(\ker[P_{21} P_{22}]) > 0$.

Assumption A1: For all $\Delta \in B\Delta$ such that $\det(I - P_{11}\Delta) = 0$, $v \in \ker(I - P_{11}\Delta)$, $v \neq 0$, implies that $P_{21}v \neq 0$.

Assumption A0 is required to ensure that the problem is nontrivial. If A0 is not satisfied then either no (w, Δ) exists, or a unique (w, Δ) describes the observed datum. In either case, the solution to Problem 4.1 is immediate. Assumption A1 essentially states that the effect of every destabilizing Δ can be observed from the output y . This is a reasonable assumption for any model used to describe a physical system.

The solution of Problem 4.1 is given by Theorem 4.3. Before stating this, a considerable amount of machinery must be introduced.

A. Formulation in Terms of Signals

Problem 4.1 will be reformulated as a series of computable conditions on the signals present in the model (see Fig. 2). For convenience, define x as the vector of unknown signals

$$x = \begin{bmatrix} v \\ w \end{bmatrix}.$$

Those vectors, x , which correspond to admissible (feasible) (w, Δ) pairs will also be referred to as admissible (feasible). Problem 4.1 can be reformulated as an optimization problem in terms of the vector x , by expressing the constraint $\Delta \in B\Delta$ as a series of norm constraints on the block components of x . Squaring these constraints results in an

optimization problem with indefinite quadratic inequality constraints and one linear equality constraint. This optimization problem is studied in detail by Smith [1], and Smith and Doyle [12], and will not be discussed further here.

For notation convenience, it will be assumed that $\dim(y) = \dim(w)$. This is without loss of generality, as P can be augmented with rows or columns of zeros. Define R_i as a projection of x onto v_i for $i = 1, \dots, m$, and the projection of x onto w for $i = m + 1$.

$$R_i = \begin{cases} \text{block row } (0_1, \dots, 0_{i-1}, I_i, 0_{i+1}, \dots, 0_m, 0_w), \\ i = 1, \dots, m \\ \text{otherwise} \\ \text{block row } (0_1, \dots, 0_m, I_w), \\ i = m + 1 \end{cases} \quad (2)$$

where $\dim(0_j) = k_i \times k_j$, $\dim(I_i) = k_i \times k_i$, $\dim(0_w) = k_i \times \dim(w)$, and $\dim(I_w) = \dim(w) \times \dim(w)$.

The constraint that

$$y = F_u(P, \Delta) \begin{bmatrix} w \\ u \end{bmatrix} = R_{m+1} P \begin{bmatrix} x \\ u \end{bmatrix} \quad (3)$$

is reformulated in terms of a subspace. Define x_0 as the solution to

$$y - P_{23}u = [P_{21} P_{22}] x$$

that is orthogonal to the kernel of $[P_{21} P_{22}]$. Note that, by Assumption A0, x_0 exists and $x_0 \neq 0$. Define a subspace \mathcal{X} by

$$\mathcal{X} = \text{span}(x_0) \oplus \ker[P_{21} P_{22}]. \quad (4)$$

As usual, the notation $B\mathcal{X}$ will be used to denote the unit ball of this subspace. For every $\hat{x} \in \mathcal{X}$ such that $\langle x_0, \hat{x} \rangle \neq 0$,

$$x = \frac{\|x_0\|^2}{\langle x_0, \hat{x} \rangle} \hat{x}$$

satisfies the equality constraint (3).

The discussion which follows will define a new system \tilde{P} . This is motivated by the consideration of a more general analysis problem involving the inclusion of known signals. Space restrictions preclude giving the full motivation for the approach. The interested reader is referred to [1].

Define \tilde{P} by

$$\tilde{P} = \begin{bmatrix} P_{11} & P_{12} \\ x_0^* / \|x_0\|^2 & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\|x_0\|^2} \begin{bmatrix} P_{13} u x_0^* \\ 0 \\ 0 \end{bmatrix} \quad (5)$$

where the zero blocks have been inserted in order to make \tilde{P} square.

B. Formulation as a μ Type Problem

It will be shown that the model validation problem can be cast into a framework reminiscent of the Fan and Tits [9] approach to the calculation of μ . The following theorem (due to Fan and Tits [9]) states that μ is equivalent to a maximization problem over the unit sphere. To maintain notational

consistency in the following, the assumed block structure has $m+1$ blocks, and R_i again defines a block projection refer to (2).

Theorem 4.2:

$$\mu(M) = \max_{\gamma, \|x\|=1} \{ \gamma \mid \|R_i x\| \gamma = \|R_i M x\|, \quad i = 1, \dots, m+1 \} \quad (6)$$

$$= \max_{\gamma, \|x\|=1} \{ \gamma \mid \|R_i x\| \gamma \leq \|R_i M x\|, \quad i = 1, \dots, m+1 \}. \quad (7)$$

Now consider the above formulation with the additional restriction that x is constrained to a subspace. Define a function $\psi(M, \mathcal{X})$ of a matrix $M \in \mathbb{C}^{n \times n}$ and a subspace $\mathcal{X} \subseteq \mathbb{C}^n$ as

$$\psi(M, \mathcal{X}) := \max_{\gamma, x \in \mathcal{X}, \|x\|=1} \{ \gamma \mid \|R_i x\| \gamma \leq \|R_i M x\|, \quad i = 1, \dots, m+1 \}. \quad (8)$$

Note that this definition is equivalent to $\mu(M)$ when $\mathcal{X} = \mathbb{C}^n$. The function $\psi(M, \mathcal{X})$ can now be used to state the main result of this section.

Theorem 4.3: Assume that Assumption A0 holds, \tilde{P} is given by (5), and \mathcal{X} is given by (4).

If $\psi(\tilde{P}, \mathcal{X}) = 0$, then no feasible (w, Δ) for Problem 4.1 exists.

If $\psi(\tilde{P}, \mathcal{X}) = \alpha > 0$, and Assumption A1 is satisfied, then there exists a feasible (w, Δ) for Problem 4.1 with

$$\|w\| \leq \alpha^{-1}, \text{ and } \|\Delta\| \leq \alpha^{-1}.$$

Furthermore, there is no feasible (w, Δ) such that both inequalities are strict.

Proof of Theorem 4.3: To prove the first part it will be shown that if a feasible (w, Δ) exists, $\psi(\tilde{P}, \mathcal{X}) > 0$. Assume a feasible (w, Δ) exists. Define

$$x = \begin{bmatrix} v \\ w \end{bmatrix}$$

where

$$v = (I - P_{11}\Delta)^{-1}(P_{12}w + P_{13}u).$$

The feasibility of Δ ensures that the above inverse is well defined. Then, for $i = 1, \dots, m$,

$$\begin{aligned} R_i x &= \Delta R_i ([P_{11}P_{12}]x + P_{13}u) \\ &= \Delta R_i \tilde{P}x \end{aligned}$$

and so

$$\|R_i x\| \leq \|\Delta\| \|R_i \tilde{P}x\|. \quad (9)$$

Furthermore,

$$\|R_{m+1}x\| = \|w\|$$

and as $\|R_{m+1}\tilde{P}x\| = 1$

$$\|R_{m+1}x\| \leq \|R_{m+1}\tilde{P}x\|. \quad (10)$$

Note that $x \in \mathcal{X}$ by assumption, and if x is scaled such that $\|x\| = 1$, the inequalities of (9) and (10) still hold. The scaled x and

$$\gamma = \min \{ \|w\|^{-1}, \|\Delta\|^{-1} \} > 0$$

satisfy the constraints of the $\psi(\tilde{P}, \mathcal{X})$ maximization (8). Consequently, $\psi(\tilde{P}, \mathcal{X}) \geq \gamma > 0$.

To prove the second part of Theorem 4.3, consider the case for $i = m+1$. $\psi(\tilde{P}, \mathcal{X}) = \alpha$ implies that there exists $x \in \mathcal{X}$, such that

$$\|R_{m+1}x\| \alpha \leq \frac{\langle x_0, x \rangle}{\|x_0\|^2}.$$

It is claimed that Assumption A1 is sufficient to guarantee that $\langle x_0, x \rangle \neq 0$. To see this assume that $\langle x_0, x \rangle = 0$, and note that $\alpha > 0$ implies that $R_{m+1}x = w = 0$. Furthermore,

$$\mathcal{X} = \text{span}(x_0) \oplus \ker [P_{21}P_{22}]$$

and so $x \in \mathcal{X}$, $\langle x_0, x \rangle = 0$ implies that $x \in \ker [P_{21}P_{22}]$. However, it has already been shown that x is of the form

$$x = \begin{bmatrix} v \\ 0 \end{bmatrix}$$

and so $v \in \ker(P_{21})$, contradicting Assumption A1. Therefore, Assumption A1 guarantees that $\langle x_0, x \rangle \neq 0$.

Now define

$$\bar{x} := \frac{\|x_0\|^2}{\langle x_0, x \rangle} x =: \begin{bmatrix} \bar{v} \\ \bar{w} \end{bmatrix}.$$

Note that

$$\frac{\langle x_0, \bar{x} \rangle}{\|\bar{x}_0\|^2} = 1$$

implying that \bar{x} satisfies the equality constraint [see (3)]. Note that \bar{x} satisfies the norm constraints of the $\psi(\tilde{P}, \mathcal{X})$ maximization. Consider the norm constraint for $i = m+1$

$$\|R_{m+1}\bar{x}\| = \|\bar{w}\| \leq \alpha^{-1} \frac{\langle x_0, \bar{x} \rangle}{\|\bar{x}_0\|^2} = \alpha^{-1}.$$

The norm constraints for $i = 1, \dots, m$ are

$$\begin{aligned} \|R_i \bar{x}\| \alpha &\leq \left\| R_i \left([P_{11}P_{12}] \bar{x} + \frac{1}{\|x_0\|^2} P_{13}u \langle x_0, \bar{x} \rangle \right) \right\| \\ &= \|R_i(P_{11}\bar{v} + P_{12}\bar{w} + P_{13}u)\| \\ &= \|\bar{z}_i\|. \end{aligned}$$

Choose

$$\Delta_i = \begin{cases} \frac{1}{\|\bar{z}_i\|^2} v_i \bar{z}_i^*, & \text{if } \bar{z}_i \neq 0 \\ 0, & \text{if } \bar{z}_i = 0. \end{cases}$$

Note that $\bar{v} = \Delta \bar{z}$ and that

$$\begin{aligned} \|\Delta\| &= \max_i \left\{ \frac{1}{\|\bar{z}_i\|^2} \|\bar{v}_i\| \|\bar{z}_i\| \right\} \\ &\leq \alpha^{-1}. \end{aligned}$$

Hence, we have constructed a feasible (w, Δ) pair with $\|w\| \leq \alpha^{-1}$ and $\|\Delta\| \leq \alpha^{-1}$.

To prove the final statement of the theorem, note that if there exists a feasible (w, Δ) with $\|\Delta\| < \alpha^{-1}$ and $\|w\| < \alpha^{-1}$, then for some $\epsilon > 0$, $\|\Delta\| \leq (\alpha + \epsilon)^{-1}$ and $\|w\| \leq (\alpha + \epsilon)^{-1}$. The construction of a scaled x , given in the first part of the proof, gives $\psi(\tilde{P}, \mathcal{X}) \geq \alpha + \epsilon$, contradicting the theorem statement that $\psi(\tilde{P}, \mathcal{X}) = \alpha$. \blacktriangleright

To find the minimum norm (w, Δ) pair, one would calculate $\psi(\tilde{P}, \mathcal{X})$. It is interesting to note that $\psi(\tilde{P}, \mathcal{X})$ can be bounded above giving a simple lower bound on the minimum norm w and Δ .

Lemma 4.4: Given $M \in \mathbb{G}^{n \times n}$ with an associated uncertainty structure Δ and a subspace \mathcal{X} , $\mathcal{X} \subseteq \mathbb{G}^n$, $\psi(M, \mathcal{X}) \leq \mu(M)$.

Proof of Lemma 4.4:

$$\begin{aligned} \psi(M, \mathcal{X}) &= \max_{\gamma, x \in \mathcal{X}, \|x\|=1} \{\gamma \mid \|R_i x\| \gamma \leq \|R_i M x\|, \\ &\quad i = 1, \dots, m+1\} \\ &\leq \max_{\gamma, \|x\|=1} \{\gamma \mid \|R_i x\| \gamma \leq \|R_i M x\|, \\ &\quad i = 1, \dots, m+1\} \\ &= \mu(M). \end{aligned}$$

This leads to the following simple bound.

Theorem 4.5: There is no feasible (w, Δ) for Problem 4.1, such that

$$\|w\| < \mu(\tilde{P})^{-1} \text{ and } \|\Delta\| < \mu(\tilde{P})^{-1}.$$

Proof of Theorem 4.5: Assume that there exists a feasible (w, Δ) such that $\|w\| < \mu(\tilde{P})^{-1}$ and $\|\Delta\| < \mu(\tilde{P})^{-1}$. Theorem 4.3 then implies that either $\|w\|$ or $\|\Delta\|$ is equal to $\psi(\tilde{P}, \mathcal{X})^{-1}$. Therefore $\psi(\tilde{P}, \mathcal{X}) > \mu(\tilde{P})$ contradicting Lemma 4.4. \blacktriangleright

Note that $\mu(\tilde{P})$ in Theorem 4.5 can be replaced by any upper bound for $\psi(\tilde{P}, \mathcal{X})$.

C. A Geometric Interpretation of $\psi(M, \mathcal{X})$

Fan and Tits [15] define and use a generalized numerical range function to calculate $\mu(M)$. It will be shown that the same approach can be used to calculate $\psi(M, \mathcal{X})$. By doing this we will also show that, like $\mu(M)$, $\psi(M, \mathcal{X})$ is equivalent to a convex optimization problem for certain block structures.

The numerical range, or field of values, of a Hermitian matrix N is defined as the set

$$\{x^* N x \mid \|x\| = 1\}. \quad (11)$$

A generalization of the numerical range is the following. Consider $m+1$ Hermitian matrices N_i , $i = 1, \dots, m+1$ of dimension $n \times n$. Define a vector valued function of x where each component of the vector v is given by

$$v_i = x^* N_i x.$$

Note that $m+1$ matrices are considered to maintain notational consistency with the definitions of $\mu(M)$ and $\psi(M, \mathcal{X})$ in Section IV-B.

Define as the generalized numerical range, the range of this function when the domain is restricted to $\|x\| = 1$

$$W(N_1, \dots, N_{m+1}) = \{v \mid v_i = x^* N_i x, \|x\| = 1\}.$$

The following lemma is proven by Fan and Tits [14] and [15].

Lemma 4.6: If $m = 1$, or $m = 2$ and $n > 2$, then $W(N_1, \dots, N_{m+1})$ is a convex set.

Now consider the application of the numerical range to the $\psi(M, \mathcal{X})$ problem. To simplify the notation assume that V is an orthonormal matrix of basis vectors for the subspace \mathcal{X} , and for all $x \in \mathcal{X}$ there exists η given by

$$x = V\eta. \quad (12)$$

Note that for all $\|\eta\| = 1$, x defined by (12) has the properties $\|x\| = 1$ and $x \in \mathcal{X}$. Define

$$N(\alpha) = \{N_1(\alpha), \dots, N_{m+1}(\alpha)\}$$

where

$$N_i(\alpha) = \alpha V^* R_i^* R_i V - V^* M^* R_i^* R_i M V.$$

Note that the domain of the numerical range function has dimension equal to $\dim(\eta)$. The numerical range of $N(\alpha)$, denoted by $W(\alpha)$, is defined as the image of the unit sphere.

$$W(\alpha) = \{v \mid v_i = \eta^* N_i(\alpha) \eta, \|\eta\| = 1\}.$$

The following lemma is immediate from Lemma 4.6 by noting that the numerical range function $W(\alpha)$ is defined by $m+1$ Hermitian matrices.

Lemma 4.7: If $m = 1$, or $m = 2$ and $\dim(\eta) > 2$, then $W(\alpha)$ is convex.

For each α , $W(\alpha)$ is a set in \mathbb{R}^{m+1} . For $m \leq 2$ this set is convex provided $\dim(\eta) > 2$. In the $\mu(M)$ case considered by Fan and Tits, there is no subspace restriction and the construction alone is sufficient to guarantee that $\dim(\eta) > 2$.

In the $\mu(M)$ case (6), the fact that there exists an x such that all constraints are met, exactly implies that $0 \in W(\alpha)$ for $\alpha = \mu(M)^2$. In the $\psi(M, \mathcal{X})$ case, the maximization problem may only have a solution with a strict inequality on one of the constraints. Consideration of the origin is no longer sufficient; the quadrant where all components of v are non-positive must also be included. Therefore, define

$$v_- = \{v \mid v_i \leq 0\}.$$

The following theorem gives an alternative representation for $\psi(M, \mathcal{X})$.

Lemma 4.8: If $\alpha \leq \psi(M, \mathcal{X})^2$, then $W(\alpha) \cap v_- \neq \emptyset$, and if $\alpha > \psi(M, \mathcal{X})^2$, then $W(\alpha) \cap v_- = \emptyset$.

Proof of Lemma 4.8: Assume that $\alpha \leq \psi(M, \mathcal{X})^2$. Assume also that \bar{x} is a solution to the $\psi(M, \mathcal{X})$ maximization. Then $\bar{x} \in \mathcal{X}$ and

$$\|R_i \bar{x}\| \alpha^{1/2} \leq \|R_i M \bar{x}\|, \quad i = 1, \dots, m+1.$$

As $\bar{x} \in \mathcal{X}$ there exists $\bar{\eta}$ given by $\bar{x} = V\bar{\eta}$. Now consider $\bar{\nu}$ where

$$\bar{\nu}_i = \bar{\eta}^* N_i(\alpha) \bar{\eta}.$$

For $i = 1, \dots, m+1$,

$$\begin{aligned} \bar{\nu}_i &= \|R_i \bar{x}\|^2 \alpha - \|R_i M \bar{x}\|^2 \\ &\leq 0 \end{aligned}$$

and so $\bar{\nu} \in \nu_-$ and $\bar{\nu} \in W(\alpha)$.

For the second part of the theorem, consider α such that there exists $\nu \in W(\alpha) \cap \nu_-$. This implies that there exists η such that

$$\eta^* N_i(\alpha) \eta \leq 0, \quad i = 1, \dots, m+1.$$

Choosing $x = V\eta$ gives

$$\|R_i x\|^2 \alpha - \|R_i M x\|^2 \leq 0, \quad i = 1, \dots, m+1$$

implying that

$$\|R_i x\| \alpha^{1/2} \leq \|R_i M x\|, \quad i = 1, \dots, m+1.$$

By construction $x \in \mathcal{X}$ and so x and $\alpha^{1/2}$ are a feasible solution to the $\psi(M, \mathcal{X})$ maximization. Therefore, $\alpha \leq \psi(M, \mathcal{X})^2$. \blacktriangleright

Corollary 4.9:

$$\psi(M, \mathcal{X}) = \inf_{\alpha \geq 0} \{ \alpha^{1/2} | W(\alpha) \cap \nu_- = \emptyset \}.$$

Fan and Tits [15] provide an algorithm to calculate $\mu(M)$, based on the numerical range approach. Their algorithm requires that one calculate the minimum distance between the set $W(\alpha)$ and the origin. To be applicable to $\psi(M, \mathcal{X})$, one must be able to calculate the minimum distance between $W(\alpha)$ and the negative quadrant ν_- . The calculation of $\psi(M, \mathcal{X})$ is the subject of ongoing research.

V. SKEWED μ PROBLEMS

The above approaches to the model validation problem have considered the means of finding a bound $\psi(\tilde{P}, \mathcal{X})$, such that there exists a feasible (w, Δ) with

$$\|w\| \leq \psi(\tilde{P}, \mathcal{X})^{-1}, \quad \text{and} \quad \|\Delta\| \leq \psi(\tilde{P}, \mathcal{X})^{-1}.$$

One can also consider the problem of finding the minimum norm w , given that Δ satisfies some *a priori* bound. The appropriate model validation problem can be stated as follows.

Problem 5.1 (Model Validation – Minimum $\|w\|$): Given a model P , with $\mu(P_{11}) < 1$, and an experimental datum (u, y) , what is the minimum $\|w\|$, such that there exists $\Delta \in B\Delta$ satisfying

$$y = F_u(P, \Delta) \begin{bmatrix} w \\ u \end{bmatrix}.$$

Note that the size of Δ is constrained, but $\|w\|$ is to be minimized. A function $\psi_s(M, \mathcal{X})$ will be defined such that there exists a feasible (w, Δ) with

$$\|w\| \leq \psi_s(\tilde{P}, \mathcal{X})^{-1}, \quad \text{and} \quad \|\Delta\| \leq 1.$$

The subscript s is used to denote the fact that the problem is

“skewed.” Section VI will show that the solution of this skewed problem leads to a solution of the model validation problem posed in Section III.

In fact, the same question can be asked for the $\mu(M)$ case. One would like to be able to calculate a bound, denoted by $\mu_s(M)$ such that (referring to the generic system in Fig. 1), for all $\|w\| \leq 1$, and all $\|\Delta\| \leq 1$, the error e is bounded by $\|e\| \leq \mu_s(M)^{-1}$.

This section will outline the results that allow both the μ and ψ problems to be skewed. For full details and proofs, the reader is referred to Smith [1]. These results are included to illustrate that it is a simple matter to adjust the relative weighting between $\|w\|$ and $\|\Delta\|$ in both the μ and ψ problems. There is nothing fundamentally new in this approach. The same results can be obtained by iteratively calculating μ (or ψ) for a scaled problem. A simple bisection technique will provide the same answer.

A. Skewed μ

A framework is set up to consider the μ problem when certain of the Δ blocks do not scale. The numerical range approach and algorithm for calculating $\mu(M)$ provided by Fan and Tits [9] can also be applied to $\mu_s(M)$.

Consider the integers $1, \dots, m+1$ to be divided into two disjoint sets: I_s and \bar{I}_s where I_s may be empty. It will now be assumed that for $i \in I_s$, $\|\Delta_i\| \leq 1$. Define a “ball” in which only certain of the Δ blocks are allowed to vary in size

$$B_\delta^s \Delta := \left\{ \Delta \mid \Delta \in \Delta, \begin{aligned} &\bar{\sigma}(\Delta_i) \leq 1, i \in I_s \\ &\bar{\sigma}(\Delta_i) \leq \delta, i \in \bar{I}_s \end{aligned} \right\}.$$

Now define “skewed μ ”, denoted μ_s , by the following:

$$\mu_s(M) := \begin{cases} 0 & \text{if no } \Delta \in B_\infty^s \Delta \text{ solves } \det(I + M\Delta) = 0 \\ \text{otherwise} \\ \min_{\Delta \in B_\infty^s \Delta} \left\{ \delta \mid \exists \Delta \in B_\delta^s \Delta \text{ such that } \det(I + M\Delta) = 0 \right\}^{-1} \end{cases}$$

Two maximization problems are now introduced

$$\hat{\mu}_s(M) := \max_{\gamma, \|x\|=1} \left\{ \gamma \mid \begin{aligned} &\|R_i x\| = \|R_i M x\|, i \in I_s \\ &\|R_i x\| \gamma = \|R_i M x\|, i \in \bar{I}_s \end{aligned} \right\}$$

and

$$\tilde{\mu}_s(M) := \max_{\gamma, \|x\|=1} \left\{ \gamma \mid \begin{aligned} &\|R_i x\| \leq \|R_i M x\|, i \in I_s \\ &\|R_i x\| \gamma \leq \|R_i M x\|, i \in \bar{I}_s \end{aligned} \right\}.$$

Note that it is not necessary to restrict x to be on the unit sphere ($\|x\| = 1$) as for any x meeting the above constraints αx , $\alpha \in (0, \infty)$ also meets the constraints. However, the restriction to the unit sphere, and the compactness of the unit sphere make it clear that the maximum is achieved.

For a finite matrix M , $\mu(M)$ is also finite. This is no longer true for $\mu_s(M)$. If the Δ_i blocks with $i \in \bar{I}_s$ do not play a role in the equation $\det(I + M\Delta) = 0$, then $\mu_s(M)$ will be infinite. In other words, there is a choice of Δ_i with

$i \in I_s$, $\|\Delta_i\| \leq 1$, such that $\det(I + M\Delta) = 0$ irrespective of the size of the Δ_i for $i \in \bar{I}_s$.

In the case where $\mu_s(M)$ is finite, the above maximizations are indeed equivalent. This can be proven in a manner similar to that used by Fan and Tits [9] for the analogous $\mu(M)$ maximization problems.

Theorem 5.2: If $\mu_s(M)$ is finite, then $\mu_s(M) = \hat{\mu}_s(M) = \tilde{\mu}_s(M)$.

B. A Geometric Interpretation

Now consider the application of the numerical range to the $\mu_s(M)$ problem. Define

$$N(\alpha) = \{N_1(\alpha), \dots, N_{m+1}(\alpha)\}$$

where $\alpha \in \mathbb{R}$, $\alpha \geq 0$, and

$$N_i(\alpha) = \begin{cases} R_i^T R_i - M^* R_i^T R_i M, & i \in I_s \\ \alpha R_i^T R_i - M^* R_i^T R_i M, & i \in \bar{I}_s. \end{cases}$$

The numerical range of $N(\alpha)$ is defined as $W(\alpha)$

$$W(\alpha) = \{v | v_i = x^* N_i(\alpha) x, \|x\| = 1\}.$$

For each α , $W(\alpha)$ is a set in \mathbb{R}^{m+1} . For $m \leq 2$, this set is convex as the above construction guarantees that $n \geq m + 1$. Define $c(\alpha)$ as the minimum distance between $W(\alpha)$ and the origin.

$$c(\alpha) = \min \{\|v\| \mid v \in W(\alpha)\}.$$

Note that as $N_i(\alpha)$ is a continuous function of α , the norm is a continuous function of its argument, and the minimization is over a compact set, $c(\alpha)$ is a continuous function of α . The function $c(\alpha)$ now provides an alternative characterization of $\mu_s(M)$ and will lead to an algorithm.

Lemma 5.3: If $\mu_s(M)$ is finite, $c(\mu_s(M)^2) = 0$, and for all $\alpha > \mu_s(M)^2$, $c(\alpha) > 0$.

Corollary 5.4: If $\mu_s(M)$ is finite,

$$\mu_s(M) = \inf_{\alpha} \{\alpha^{1/2} \mid 0 \notin W(\beta) \text{ for all } \beta > \alpha\}.$$

C. An Algorithm for $\mu_s(M)$

This is the same algorithm as that presented by Fan and Tits [15] for the calculation of $\mu(M)$. The proof of convergence is also identical. This algorithm actually calculates $\hat{\mu}_s(M)$. The assumption that $\mu_s(M)$ is finite, is disguised in step i.

Algorithm 5.5 ($\mu_s(M)$):

- i) Set $\alpha_0 \geq \mu_s(M)^2$.
- ii) $\alpha_{j+1} = \alpha_j - c(\alpha_j)$.
- iii) $j = j + 1$. Go to step ii).

Lemma 5.6: The sequence $\{\alpha_j\}$ generated by Algorithm 5.5 is monotonic nonincreasing and

$$\lim_{j \rightarrow \infty} \alpha_j = \mu_s(M)^2.$$

It should be noted that for $\alpha < \mu_s(M)^2$, it is not necessarily true that $0 \in W(\alpha)$. This may make the choice of α_0 difficult. In the case where $I_s = \emptyset$, $\mu_s(M) = \mu(M)$, and the obvious choice is to use the upper bound for $\mu(M)$: $\alpha_0 = \bar{\sigma}(M)^2$.

When $\mu_s(M)$ is not finite, it is not possible to choose α_0 . If $\alpha_0 < \mu_s(M)^2$, the algorithm may converge to an incorrect finite value.

Critical to Algorithm 5.5 is the calculation of $c(\alpha)$. In general this is still an unsolved problem. It is, however, easy to calculate the minimum point of a convex set. Define $c'(\alpha)$ by

$$c'(\alpha) = \min \{\|v\| \mid v \in \text{co}[W(\alpha)]\}.$$

Gilbert [16] provides an algorithm that will calculate $c'(\alpha)$. This is discussed by Doyle [17] and Packard [11], both of whom provide a proof of convergence.

Note that $c'(\alpha) \leq c(\alpha)$ and that $W(\alpha)$ being convex is sufficient, but not necessary, to guarantee that $c'(\alpha) = c(\alpha)$. In practice, one can replace $c(\alpha)$ in Algorithm 5.5 by $c'(\alpha)$. This converges to an upper bound for $\mu_s(M)^2$. In the case of three or fewer blocks, Lemma 4.6 illustrates that it will converge to $\mu_s(M)^2$.

This supports Doyle's [17] result that $\mu(M)$ can easily be calculated for three or fewer complex blocks ($m \leq 2$). The fact that this is also true for $\mu_s(M)$ is hardly surprising as $\mu_s(M)$ can be iteratively calculated from $\mu(M)$.

D. A Skewed $\mu(M, \mathcal{X})$ Problem

As in the $\mu_s(M)$ case, one can define a skewed version of $\psi(M, \mathcal{X})$, denoted here by $\psi_s(M, \mathcal{X})$.

$$\psi_s(M, \mathcal{X}) := \begin{cases} 0 & \text{if } \{x \mid x \in \mathcal{X}, \|x\| = 1, \\ & \|R_i x\| \leq \|R_i M x\|, i \in I_s\} = \emptyset \\ \text{otherwise} \\ \max_{\gamma, x \in \mathcal{X}, \|x\|=1} & \left\{ \gamma \mid \begin{aligned} & \|R_i x\| \leq \|R_i M x\|, i \in I_s \\ & \|R_i x\| \gamma \leq \|R_i M x\|, i \in \bar{I}_s \end{aligned} \right\} \end{cases} \quad (13)$$

E. Application to the Minimum $\|w\|$ Model Validation Problem

The function $\psi_s(M, \mathcal{X})$ can be applied to Problem 5.1. Define $I_s = 1, \dots, m$ and $\bar{I}_s = m + 1$. In other words, the scaling on the perturbation blocks is held fixed. Again, \tilde{P} is defined by (5), and \mathcal{X} by (4). The following theorem gives the bound on the smallest feasible $\|w\|$.

Theorem 5.7: Assume that Assumption A0 holds, \tilde{P} is given by (5), and \mathcal{X} is given by (4).

If $\psi_s(\tilde{P}, \mathcal{X}) = 0$, then no feasible (w, Δ) for Problem 5.1 exists.

If $\psi_s(\tilde{P}, \mathcal{X}) = \alpha > 0$, then there exists a feasible (w, Δ) for Problem 5.1 with

$$\|w\| = \alpha^{-1}, \quad \text{and } \|\Delta\| \leq 1.$$

Furthermore, there is no feasible (w, Δ) such that $\Delta \in B\Delta$, and $\|w\| < \alpha^{-1}$.

The assumption of Problem 5.1 that $\mu(P_{11}) < 1$, and the restriction of Δ to $\Delta \in B\Delta$ obviates the need for inclusion of Assumption A1 in the above.

F. A Numerical Range Function

The $\psi_s(M, \mathcal{X})$ calculation problem (13) is considered for a general matrix M and a subspace \mathcal{X} . It is again assumed in the notation used in this section that the general problem has $m+1$ norm constraints; in other words $I_s \cup \bar{I}_s = \{1, \dots, m+1\}$. Define

$$N_s(\alpha) = \{N_{s1}(\alpha), \dots, N_{s(m+1)}(\alpha)\}$$

where

$$N_{si}(\alpha) = \begin{cases} V^* R_i^T R_i V - V^* M^* R_i^T R_i M V, & i \in I_s \\ \alpha V^* R_i^T R_i V - V^* M^* R_i^T R_i M V, & i \in \bar{I}_s. \end{cases}$$

The numerical range of $N_s(\alpha)$ is defined as $W_s(\alpha)$

$$W_s(\alpha) = \{\nu | \nu_i = \eta^* N_{si}(\alpha) \eta, \|\eta\| = 1\}. \quad (14)$$

Note that the convexity properties (refer to Lemma 4.7) of $W_s(\alpha)$ are the same as $W(\alpha)$.

A lemma, analogous to Lemma 4.8 for $\psi(M, \mathcal{X})$, and Lemma 5.3 for $\mu_s(M)$ illustrates a direction that one might take in deriving an algorithm for $\psi_s(M, \mathcal{X})$.

Lemma 5.8: If $\alpha \leq \psi_s(M, \mathcal{X})^2$, then $W_s(\alpha) \cap \nu_- \neq \emptyset$, and if $\alpha > \psi_s(M, \mathcal{X})^2$, then $W_s(\alpha) \cap \nu_- = \emptyset$.

Corollary 5.9:

$$\psi_s(M, \mathcal{X}) = \inf_{\alpha \geq 0} \{\alpha^{1/2} | W_s(\alpha) \cap \nu_- \neq \emptyset\}.$$

As is the case for $\psi(M, \mathcal{X})$, using Corollary 5.9 as the basis for an algorithm for $\psi_s(M, \mathcal{X})$ relies upon being able to calculate the minimum distance between $W_s(\alpha)$ and ν_- . As expected, $\psi_s(M, \mathcal{X})$ can easily be bounded.

Lemma 5.10:

$$\psi_s(M, \mathcal{X}) \leq \mu_s(M).$$

G. Approaches to the Calculation of ψ_s

The framework discussed above suggests that one should be able to calculate ψ_s using an algorithm based on the Fan and Tits approach to the calculation μ . This is the subject of ongoing research.

The function ψ_s is defined in terms of an optimization problem. Another calculation approach is to apply standard nonlinear optimization techniques to the definition of ψ_s given in (13). For simple block structures (two or less blocks with sufficient freedom in kernel of $[P_{21} P_{22}]$) the optimization is convex.

Upper and lower bounds can readily be calculated for μ . The lower bound is calculated via a power iteration. Applying this to the ψ_s is the subject of current research. The

upper bound for μ can be calculated by solving a linear matrix inequality (LMI) problem. The same is true of the ψ_s case although the formulation of the LMI is now slightly more complicated. Newlin and Smith [18] discuss these calculation approaches in more detail.

VI. MODEL VALIDATION FOR μ/H_∞ REVISITED

The linear, time-invariant model validation problem posed in Problem 3.1 can now be studied in terms of the above constant matrix model validation problems. The following theorem addresses Problem 3.1.

Assume that A0 holds at each frequency. If this is not the case then either no feasible $(w(j\omega), \Delta(j\omega))$ exists, or a unique $w(j\omega)$ solves the constant matrix problem. If such a $w(j\omega)$ exists then $\psi_s(\tilde{P}(j\omega), \mathcal{X}(j\omega))^2$ can be replaced by $\|w(j\omega)\|^2$ in the following theorem.

Theorem 6.1 (Model Validation): Assume that, at each frequency, Assumption A0 holds, $\tilde{P}(j\omega)$ is given by (5), and $\mathcal{X}(j\omega)$ is given by (4).

If

$$\int_{-\infty}^{\infty} \frac{1}{\psi_s(\tilde{P}(j\omega), \mathcal{X}(j\omega))^2} d\omega > 1$$

then there is no $w \in BL_2$ and $\Delta \in B\Delta$, such that

$$y = F_u(P, \Delta) \begin{bmatrix} w \\ u \end{bmatrix}.$$

Proof of Theorem 6.1: Note that, by definition, $\psi_s(\tilde{P}(j\omega), \mathcal{X}(j\omega))^{-1} \leq \|w(j\omega)\|$, for any $w(j\omega)$ satisfying

$$y = F_u(P, \Delta) \begin{bmatrix} w \\ u \end{bmatrix}$$

for some $\Delta(j\omega) \in B\Delta$. Consequently,

$$\int_{-\infty}^{\infty} \psi_s(\tilde{P}(j\omega), \mathcal{X}(j\omega))^{-2} d\omega \leq \|w\|^2$$

for any (w, Δ) satisfying

$$y = F_u(P, \Delta) \begin{bmatrix} w \\ u \end{bmatrix}.$$

To apply this theorem, one typically selects a frequency grid, calculates $\psi_s(\tilde{P}(j\omega), \mathcal{X}(j\omega))$ (or equivalently, the minimum $\|w(j\omega)\|$ satisfying the constraints of the problem), and approximates the integration of Theorem 6.1 with a summation.

Note that the ψ_s problem, rather than the ψ problem, is applicable here. In order to be assured of obtaining a bound for the minimum $\|w\|$ signal, $w \in L_2$, one must bound the minimum $\|w(j\omega)\|$ at each frequency.

It is also worth noting that Theorem 6.1 does not necessarily apply in the opposite direction. By solving a series of minimum $\|w(j\omega)\|$ problems one obtains complex valued $w(j\omega)$ and $\Delta(j\omega)$ satisfying the constraints at each frequency. However this does not necessarily correspond to a causal Δ .

The lower bound obtained from Theorem 6.1 is appropriate for the model validation problem as it still yields a

necessary condition for an uncertain model to describe a physical system.

VII. CONCLUSIONS

The model validation problem has been formulated for a generic robust control structure. The significance of the formulation given here is that the assumptions used are exactly the same as those required for the H_∞/μ synthesis methodology. It should be noted that using these assumptions imposes significant constraints on the type of experiment that can be considered.

Two constant matrix model validation problems are studied. The first is, "what is the smallest $\|w\|$ and $\|\Delta\|$ accounting for the datum?" A function $\psi(M, \mathcal{X})$ has been defined and shown to give the solution to the above problem when applied to a particular interconnection structure. A geometric interpretation allows one to see that $\psi(M, \mathcal{X})$ is very similar to $\mu(M)$. In particular, the convexity properties of $W(\alpha)$ (which lead to the upper bound being an equality for certain block structures) apply also to the $\psi(M, \mathcal{X})$ case.

Of perhaps more engineering significance is the second constant matrix problem: "What is the smallest $\|w\|$ such that w and a $\Delta \in B\Delta$, account for the datum?" Skewed problems have been introduced for both $\mu(M)$ and $\psi(M, \mathcal{X})$ and shown to lead to the solution of the minimum $\|w\|$ model validation problem. Note that for both $\mu_s(M)$ and $\psi_s(M, \mathcal{X})$, the same result can be calculated iteratively with a scaled $\mu(M)$, or $\psi(M, \mathcal{X})$ problem. A number of approaches to calculating $\psi_s(M, \mathcal{X})$ have been outlined. It is no surprise that these are very similar to the approaches used for the calculation of μ .

The constant matrix model validation results are used to give a necessary test for a robust control model (in the H_∞/μ framework) to account for an experimental datum.

Several issues immediately arise as future research directions. Algorithms to calculate (or bound) $\psi(M, \mathcal{X})$, and possibly $\psi_s(M, \mathcal{X})$, require more work. Note that also that only full complex block structures have been studied. To be applicable to all possible robust control models, the results should be extended to more general structures. There are now many applications for which μ synthesis controllers have been designed and the models used for the design can now be studied with the model validation approach.

The formal approach to model validation given here can also be applied to other design methodologies. Research on the application to h_∞ (discrete time) and l_1 methodologies is in progress.

On a more long-term basis, it appears possible that this approach can be used to study the general problem of identification of robust control models. Refer to Smith and Doyle [19] for further discussion on this issue.

ACKNOWLEDGMENT

The authors would like to thank M. Newlin, H. Stalford, A. Packard, K. Zhou, K. Glover, and M. A. Dahleh for many helpful discussions and reading various versions of this work.

REFERENCES

- [1] R. S. Smith, "Model validation for uncertain systems," Ph.D. dissertation, California Inst. Technol., Pasadena, CA, 1990.
- [2] K. Poolla, Private communication, 1991.
- [3] L. Ljung, *System Identification, Theory for the User* (Information and System Science Series). Englewood Cliffs, NJ: Prentice-Hall, 1987.
- [4] A. Helmicki, C. Jacobson, and C. Nett, " H_∞ identification of stable LSI systems: A scheme with direct application to controller design," in *Proc. Amer. Contr. Conf.*, 1989, pp. 1428-1434.
- [5] —, "Control oriented system identification: A worst-case/deterministic approach in H_∞ ," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 1163-1176, 1991.
- [6] G. Gu and P. P. Khargonekar, "Linear and nonlinear algorithms for identification in H^∞ with error bounds," in *Proc. Amer. Contr. Conf.*, 1991, pp. 64-69.
- [7] P. Mäkilä, "Laguerre methods and H^∞ identification of continuous-time systems," *Int. J. Contr.*, vol. 53, pp. 689-707, 1991.
- [8] P. Mäkilä and J. Partington, "Robust approximation and identification in H^∞ ," in *Proc. Amer. Contr. Conf.*, 1991, pp. 70-76.
- [9] M. K. H. Fan and A. L. Tits, "Characterization and efficient computation of the structured singular value," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 734-743, 1986.
- [10] J. Doyle, "Structured uncertainty in control system design," in *Proc. IEEE Conf. Decision Contr.*, 1985, pp. 260-265.
- [11] A. K. Packard, "What's new with μ ," Ph.D. dissertation, Univ. California, Berkeley, CA, 1988.
- [12] R. S. Smith and J. Doyle, "Model invalidation—A connection between robust control and identification," in *Proc. Amer. Contr. Conf.*, 1989, pp. 1435-1440.
- [13] M. K. H. Fan and A. L. Tits, "Geometric aspects in the computation of the structured singular value," in *Proc. Amer. Contr. Conf.*, 1986, pp. 437-441.
- [14] —, "On the generalized numerical range," *Linear and Multilinear Algebra*, vol. 21, pp. 313-320, 1987.
- [15] —, " m -form numerical range and the computation of the structured singular value," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 284-289, 1988.
- [16] E. Gilbert, "An iterative method for computing the minimum of a quadratic form on a set," *SIAM J. Contr.*, vol. 4, pp. 61-80, 1966.
- [17] J. Doyle, "Analysis of feedback systems with structured uncertainties," *IEE Proc. Part D*, vol. 133, pp. 45-56, Mar. 1982.
- [18] M. Newlin and R. S. Smith, "Model validation and generalized μ ," in *Proc. IEEE Conf. Decision Contr.*, 1991, pp. 1257-1258.
- [19] R. S. Smith and J. C. Doyle, "Towards a methodology for robust parameter identification," in *Proc. Amer. Contr. Conf.*, 1990, pp. 2394-2399.



Roy S. Smith (S'79-M'81-S'85-M'89) received the B.E. (Hons.) and the M.E. degrees from the University of Canterbury, New Zealand, in 1980 and 1981, respectively, and the M.S.E.E. and Ph.D. degrees from the California Institute of Technology, Pasadena, in 1986 and 1990, respectively.

He has worked, in New Zealand, for the Energy Research and Development Committee, Worley Downey Mandeno, and the Institute of Nuclear Science on the control and instrumentation of automotive engines, boiler installations, linear accelerators and mass spectrometers, and, in California, for Olson Engineering on vehicle emissions and fuel control systems. In 1989 he held a joint position, Lecturer at Caltech, and working for the Jet Propulsion Laboratory on the control of flexible space structures. In 1990 he joined the University of California, Santa Barbara, as an Assistant Professor in the Department of Electrical and Computer Engineering. His research interests include robust control, identification with uncertain models, and the application of these to automotive systems, process control, and flexible structures.

Dr. Smith is a member of SIAM, AIAA, and NZAC.



John C. Doyle received the B.S. and M.S. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1977, and the Ph.D. degree in mathematics from the University of California, Berkeley, in 1984.

He is a Professor of Electrical Engineering at Caltech, Pasadena, and has been a consultant to Honeywell Systems and Research Center since 1976. His theoretical research interests include modeling and control of uncertain and nonlinear systems, matrix perturbation problems, operator theoretic methods, and μ . His theoretical work has been applied throughout the aerospace industry and is gaining acceptance in the process control

industry. His current application interests include flexible structures, chemical process control, flight control, and control of unsteady fluid flow and combustion. Additional academic interests include the impact of control on system design, the role of neoteny in personal and social evolution, modeling and control of acute and chronic human response to exercise, and feminist critical theory, especially in the philosophy of science.

Dr. Doyle is the recipient of the Hickernell Award, the Eckman Award, the IEEE Control Systems Society Centennial Outstanding Young Engineer Award, and the Bernard Friedman Award. He is an NSF Presidential Young Investigator, or ONR Young Investigator, and has coauthored two TRANSACTIONS Best Paper Award winners, one of which won the IEEE Baker Prize.